



Improving Writing Skill of Non-English Department Students Using AI

Rachmat Ari Wibowo
Universitas Ivet
Rachmat.ari@unisvet.ac.id

Abstrak

The integration of artificial intelligence (AI) technology into education has become increasingly important in supporting the writing proficiency of non-English major students in *English for Specific Purposes* (ESP) courses. This study investigates the effectiveness of ChatGPT in improving the clarity, coherence, and grammatical accuracy of students' writing drafts. Employing a quasi-experimental design, the study examined differences in learning outcomes between a ChatGPT-assisted class (experimental group) and a traditionally taught class (control group). Data were collected through writing draft assessments, participant surveys, and interviews. The findings indicate that the experimental group demonstrated a significant improvement in writing performance, particularly in organizing and presenting ideas clearly and coherently. Furthermore, the personalized guidance provided by ChatGPT contributed to students' learning autonomy and self-confidence. These findings highlight the potential of ChatGPT as a supportive tool for language learning while emphasizing the continued role of lecturers in validating recommendations generated by AI systems. Future research is encouraged to explore AI-assisted writing in more specialized disciplinary contexts and to investigate its long-term effects on students' learning retention and writing development.

Kata Kunci: Artificial Intelligence in Education, ChatGPT-Assisted Writing, *English for Specific Purposes* (ESP)

Introduction

English writing proficiency constitutes an important academic competency in higher education, particularly for students enrolled in non-English departments who take English for Specific Purposes (ESP) courses. Despite its importance, many EFL learners continue to experience difficulties in producing texts that communicate ideas clearly, demonstrate logical organization, and employ accurate language. These challenges involve not only

grammatical and lexical problems but also difficulties in developing coherent arguments and presenting ideas in an effective manner. In response to these challenges, contemporary language pedagogy has increasingly explored the integration of conventional classroom instruction with Artificial Intelligence (AI)-based technologies, particularly generative AI applications such as ChatGPT.



Artificial Intelligence has introduced new possibilities for supporting the development of academic writing. In this study, AI refers to generative systems powered by large language models (LLMs) that are capable of processing, understanding, and generating human-like language. ChatGPT serves as the primary AI-based tool examined in the present study. During the writing process, ChatGPT can assist learners by suggesting alternative expressions, reorganizing ideas, identifying potential linguistic problems, and offering recommendations for improving their drafts. Within this context, feedback is understood as information provided to learners regarding the quality of their performance that can subsequently guide improvement and revision. More specifically, personalized feedback refers to immediate, specific, and learner-oriented responses that address the characteristics of an individual student's writing. Through repeated interactions with ChatGPT under appropriate instructor supervision, students can receive feedback directly related to their own drafts rather than relying solely on generalized examples or teacher-provided corrections.

The integration of AI into language learning has consequently contributed to changes in the landscape of academic writing and technology-enhanced learning, including the application of digital technologies in learning activities (Wahyuni et al., 2020). Generative AI tools provide learners with opportunities to obtain immediate responses to their writing, allowing them to identify and revise problems without necessarily having to wait for direct feedback from their

instructors (Kasneci et al., 2023). Such immediate interaction may encourage students to engage in continuous drafting and revision and, consequently, become more actively involved in their own learning processes.

To establish clear and measurable writing outcomes, the present study conceptualizes writing performance through three major dimensions: clarity, coherence, and accuracy. These constructs are operationalized using an analytic writing rubric. Clarity refers to the extent to which the main ideas and supporting information are communicated explicitly and effectively. It is assessed based on the clarity of the thesis or central argument, the relevance of supporting details, and the precision of word choice. Coherence refers to the logical and meaningful relationship among ideas within a text. It is evaluated through the overall organization of the text, unity within paragraphs, and the appropriate use of cohesive devices. Accuracy, meanwhile, concerns the correctness of linguistic forms, including grammatical structures, vocabulary, and writing mechanics. It is assessed through indicators related to syntax, lexical choice, grammar, spelling, punctuation, and other mechanical features. Each dimension is evaluated using a 0–4 scale, while error density, calculated as the number of errors per 100 words, is incorporated as an additional indicator to strengthen the measurement of linguistic accuracy.

These dimensions are grounded in established perspectives on writing development and are therefore considered



central outcome variables in the present study. Clarity and coherence primarily represent higher-order cognitive and rhetorical abilities because they require learners to select, organize, connect, and communicate ideas effectively. Accuracy, in contrast, reflects learners' control over linguistic forms and conventions. Considered together, these three dimensions represent both the cognitive and linguistic demands involved in ESP writing. Although previous research on AI-assisted writing has frequently emphasized improvements in grammatical and linguistic accuracy, comparatively less attention has been given to the development of clarity and coherence. This imbalance indicates a conceptual gap in understanding the broader contribution of AI to writing performance. By explicitly distinguishing these dimensions, the present study seeks to provide a more comprehensive examination of how ChatGPT may contribute differently to various aspects of writing among non-English majors.

As an interactive writing assistant, ChatGPT can potentially help learners improve the organization and development of their texts by offering suggestions related to coherence, cohesion, and the logical progression of ideas (Zhai, 2023). In addition, AI-generated feedback may support improvements in linguistic accuracy, particularly in grammar, vocabulary selection, and syntactic construction, all of which constitute important elements of effective academic communication (Dwivedi et al., 2023). These capabilities are especially relevant to ESP learning contexts, in which students from non-English departments may

encounter considerable difficulties in expressing discipline-related ideas clearly and accurately in English. Nevertheless, the use of AI should not be understood as a replacement for human instruction. Previous scholars have emphasized the importance of teacher involvement and critical human mediation to ensure that AI-generated feedback is appropriate, pedagogically relevant, and consistent with the linguistic and disciplinary conventions of academic writing (Smutny & Schreiberova, 2020).

Recent empirical research has provided growing evidence of the potential contribution of AI to the development of learners' writing abilities across different educational settings. Studies have indicated that AI-supported writing can contribute not only to improvements in linguistic accuracy but also to broader aspects of writing development (Azennoud, 2024; Pratama & Sulistiyo, 2024; Hyland, 2019). For example, Azennoud (2024) reported that Moroccan EFL learners who incorporated AI tools into their writing practices demonstrated improvements in grammatical accuracy and syntactic complexity. This finding suggests that AI-supported writing may facilitate development beyond simple error correction by encouraging learners to construct more varied and sophisticated sentences. Similarly, a systematic review conducted by Pratama and Sulistiyo (2024) found that AI-assisted feedback can increase students' motivation, encourage repeated revision, and contribute to measurable improvements in writing performance. The emphasis on continuous revision is consistent with the process-oriented perspective of writing proposed by



Hyland (2019), which views writing as a recursive process involving planning, drafting, revising, and refining ideas rather than as a single, completed product.

The positive potential of AI-assisted writing has also been documented in the Indonesian educational context. Erito (2023), for instance, found that students perceived AI as a useful source of support that could reduce anxiety associated with writing and increase their confidence in expressing ideas in English. This affective aspect is particularly significant because lower levels of language anxiety have been associated with greater confidence and willingness to participate in language-learning and communication activities (Horwitz, 2016). In a similar vein, Fitriani and Fakhrudin (2024) reported that students appreciated the practical functions of AI applications, particularly their ability to check grammatical errors, generate ideas, and provide initial feedback on written work. These functions may contribute to a more efficient and independent writing process by enabling students to identify potential problems and make revisions before submitting their work to instructors.

The findings from these studies suggest that learners increasingly use AI not merely as passive recipients of automated corrections but as active participants who strategically employ technological resources to address difficulties encountered during the writing process. Such strategic use of AI can potentially strengthen learner autonomy by allowing students to seek clarification, evaluate alternative expressions, revise their drafts, and monitor their own progress. At the

same time, effective use of AI requires students to exercise judgment regarding the quality, relevance, and appropriateness of the feedback generated by the system.

Overall, the existing literature demonstrates considerable pedagogical potential for integrating AI into English writing instruction. AI-based tools such as ChatGPT may support both cognitive and affective dimensions of learning by facilitating improvements in accuracy, coherence, organization, and linguistic complexity while simultaneously contributing to students' motivation, confidence, and reduced writing anxiety. These benefits suggest that AI can become a valuable resource for promoting learner autonomy and supporting iterative writing development. However, the effective implementation of AI in education requires a balanced approach in which technological assistance is integrated with teacher guidance. Students need to be encouraged to critically evaluate AI-generated suggestions rather than accepting them uncritically. Teacher mediation remains essential for helping learners determine whether suggested revisions are accurate, contextually appropriate, and consistent with academic and disciplinary expectations. Such an approach can also foster students' metacognitive awareness by encouraging them to reflect on their writing decisions and develop more effective strategies for planning, monitoring, and revising their texts (Kasneci et al., 2023; Zawacki-Richter et al., 2019). Therefore, examining the contribution of ChatGPT to clarity, coherence, and accuracy may provide a more comprehensive



understanding of the role of generative AI in supporting writing development among non-English majors in ESP contexts.

Traditional instruction continues to play a fundamental role in academic writing because it provides opportunities for teachers to offer individualized guidance, demonstrate disciplinary conventions, and develop students' higher-order cognitive abilities, including argumentation, critical thinking, and analytical skills. Direct interaction with instructors enables students to gain a deeper understanding of syntactic structures, academic conventions, and rhetorical practices that are necessary for producing well-organized and logically developed texts. Harahap et al. (2023), for example, demonstrated that reading habits, critical thinking abilities, and mastery of syntactic structures collectively exerted a significant influence on students' scientific writing proficiency. Similarly, Tanjung et al. (2025) found that direct instructional strategies could strengthen students' critical thinking skills by creating structured learning environments in which teachers actively modeled, guided, and scaffolded processes of analysis and evaluation. Furthermore, Yuliana et al. (2020) reported that students who participated in a combination of teacher guidance and inquiry-based learning through the Argument Based Science Inquiry (ABSI) model demonstrated stronger argumentation and critical thinking abilities than students who participated exclusively in inquiry-based learning. Collectively, these findings demonstrate that although student-centered and technology-enhanced approaches offer important pedagogical benefits, teacher-led

instruction remains crucial for developing disciplinary literacy, academic writing competence, and higher-order cognitive skills.

Another important strength of traditional instruction lies in the role of instructor feedback. Research has established that teacher feedback can positively contribute to student learning when it is timely, specific, constructive, and closely aligned with instructional objectives (Hyland & Hyland, 2019). Through direct feedback, instructors can identify individual weaknesses and provide explanations that help learners understand not only what needs to be corrected but also why a particular revision is necessary. However, the effectiveness of teacher feedback may be constrained by several practical challenges. One major limitation concerns the time required to provide responses. Feedback is frequently delivered after a substantial delay, reducing students' opportunities to immediately incorporate recommendations into subsequent revisions. Although students generally value prompt feedback, instructors often manage considerable teaching and administrative responsibilities, which can result in delayed responses (Ferris, 2014).

Accordingly, this study investigates the effectiveness of ChatGPT in improving the academic writing performance of non-English majors. Specifically, it compares the writing outcomes of students who receive ChatGPT-assisted instruction with those of students who participate in traditional teacher-led instruction. In addition to examining writing performance, the study explores students' perceptions of ChatGPT in



relation to their confidence and autonomy throughout the writing process. Based on the identified theoretical and empirical gaps, the study addresses the following research questions:

1. To what extent does ChatGPT-assisted instruction improve students' clarity, coherence, and accuracy compared with traditional teaching methods?
2. How do students perceive the use of ChatGPT in relation to their confidence and autonomy in academic writing?

By addressing these questions, the study aims to contribute empirical evidence to the expanding body of scholarship on AI-supported language learning and academic writing. More specifically, it seeks to provide a comprehensive understanding of the extent to which generative AI can support different dimensions of writing performance and to offer pedagogical implications for the responsible integration of ChatGPT into ESP instruction for non-English majors.

Literature Review

Research into the application of Artificial Intelligence (AI) in education, particularly in the development of academic writing, has grown considerably in recent years. The expanding body of literature, however, indicates that the contribution of AI to writing development cannot be understood solely in terms of its effectiveness. Rather, its use involves a complex interaction among linguistic improvement, cognitive engagement, affective responses, and pedagogical practices. Existing studies can

generally be grouped into several interconnected areas. The first focuses on the capacity of AI to improve linguistic and textual features, while the second examines its potential to support higher-order writing abilities. A third body of research explores affective outcomes such as motivation, anxiety, and learner autonomy, whereas another emphasizes the importance of teacher involvement and appropriate pedagogical integration.

AI and the Improvement of Textual Features

A considerable body of research has demonstrated the potential of AI-based technologies, particularly large language models (LLMs), to enhance various linguistic dimensions of writing. ChatGPT, for example, can provide rapid feedback concerning grammatical structures, vocabulary selection, sentence construction, and other linguistic features (Kasneci et al., 2023; Zhai, 2023). Such immediate feedback enables learners to identify potential errors and make revisions during the drafting process. Empirical evidence also supports these observations. Azennoud (2024), for instance, reported improvements in learners' grammatical accuracy and syntactic complexity following the integration of AI tools into writing activities.

Despite these promising findings, a substantial proportion of existing research has concentrated on relatively observable and measurable linguistic features. Improvements in grammar, vocabulary, and sentence-level construction are often emphasized, whereas less attention has been



devoted to broader textual qualities. In particular, the influence of AI on the clarity of ideas, logical organization, and coherence of arguments remains comparatively underexplored. This tendency toward micro-level linguistic outcomes indicates the need for research that extends beyond grammatical correctness and examines how AI contributes to the development of more complex and higher-order dimensions of academic writing.

AI, Coherence, Clarity, and Higher-Order Thinking

Beyond sentence-level accuracy, some studies have investigated the potential of AI to support the organization and development of ideas within written texts. AI-generated suggestions may assist learners in arranging information, establishing relationships among ideas, and improving the structural flow of their writing (Zhai, 2023; Generative AI-Assisted Feedback, 2025). By interacting with AI during the drafting and revision stages, learners may obtain alternative ways of expressing ideas and organizing their arguments, which could potentially contribute to greater textual coherence.

Nevertheless, the ability of AI to support deeper rhetorical and cognitive processes remains subject to debate. Critical examinations have suggested that generative AI systems may have limitations in understanding contextual nuances, disciplinary conventions, and domain-specific reasoning (Razi, 2023; Chan & Tsi, 2023). Although AI can identify structural inconsistencies or suggest connections

between textual elements, it may not fully understand the conceptual significance or disciplinary implications of an argument. Consequently, AI-generated suggestions may improve surface-level organization without necessarily strengthening the underlying quality of reasoning or the clarity of the writer's intended meaning.

These limitations are particularly important when writing is viewed as a meaning-making and problem-solving process rather than simply an exercise in linguistic production. From a writing-process perspective, effective writing involves generating ideas, organizing information, constructing arguments, evaluating alternatives, and repeatedly revising meaning. These processes require cognitive engagement and contextual judgment that cannot necessarily be delegated to automated systems. Therefore, while AI may serve as a useful resource for supporting coherence and organization, human judgment remains essential for developing substantive arguments, interpreting contextual meanings, and ensuring that written ideas are logically and persuasively communicated.

Affective Dimensions: Motivation, Anxiety, and Learner Autonomy

In addition to linguistic and cognitive outcomes, an increasing number of studies have examined the affective consequences of integrating AI into writing instruction. Research by Wang et al. (2023), Erito (2023), and Apriani et al. (2024) indicates that AI-supported learning may reduce students' anxiety and increase their motivation to engage in writing activities. One possible explanation is the immediate and relatively



non-judgmental nature of AI-generated feedback. Unlike conventional feedback, which may require students to wait for teacher responses, AI allows learners to obtain suggestions and clarification at the point when they encounter difficulties.

The affective benefits of AI are also related to the concept of feedback literacy, which emphasizes learners' ability to understand, evaluate, and effectively apply feedback to improve their performance (Oktarin et al., 2024). From this perspective, the value of AI feedback does not depend solely on the amount of information provided but also on students' ability to interpret the suggestions and make appropriate decisions regarding revision. AI may therefore provide additional opportunities for learners to engage actively with feedback and take greater responsibility for their writing development.

However, several limitations should be considered when interpreting these findings. Much of the existing research on motivation, anxiety, and autonomy relies heavily on students' self-reported perceptions and involves relatively short intervention periods. Such methodological characteristics make it difficult to determine whether positive affective responses are sustained over time or whether increased motivation and reduced anxiety ultimately result in measurable and lasting improvements in writing competence. Further research is therefore needed to establish stronger connections between affective benefits and actual long-term writing development.

The Role of Educators and Pedagogical Integration

The effectiveness of AI in educational settings depends not only on the capabilities of the technology but also on how it is incorporated into instructional practices. Rather than functioning as an independent substitute for teachers, AI appears to be most beneficial when combined with appropriate teacher mediation. Educators remain responsible for helping students understand the limitations of AI-generated content, evaluate the quality of automated feedback, and apply suggestions in ways that are consistent with academic and disciplinary expectations.

Recent research emphasizes the importance of teacher involvement in AI-enhanced learning environments. Liu and Fan (2024), for example, found that teacher support significantly influences learners' willingness to communicate in technology-mediated environments. Their findings reinforce the argument that effective AI integration requires teachers to create supportive learning conditions in which students can use technology while continuing to receive appropriate human guidance. Similarly, Willian et al. (2024) highlight the importance of educators' adaptability and resourcefulness in combining AI-based tools with conventional teaching strategies. Such integration can create more dynamic learning environments while maintaining the instructional and interpersonal functions of teachers.

The need for teacher mediation is particularly important because excessive



dependence on AI may encourage learners to accept automated suggestions without sufficient critical evaluation. Wang and Han (2022) caution that inadequate guidance may result in overreliance on AI or inappropriate use of its outputs. Students may assume that AI-generated responses are automatically accurate, even though such systems can produce information that is contextually inappropriate, conceptually weak, or inconsistent with disciplinary conventions. Teacher intervention is therefore necessary to develop students' critical awareness and ensure that AI is used as a learning resource rather than as an unquestioned authority.

In response to these concerns, several scholars have proposed blended instructional models that combine AI-supported activities with teacher-led instruction (Datskiv et al., 2024; Sari & Han, 2024). Such approaches recognize the complementary strengths of both human instruction and AI technology. AI can provide immediate assistance, facilitate repeated practice, and support individualized revision, while teachers can provide contextual interpretation, disciplinary expertise, critical evaluation, and higher-order guidance. Nevertheless, empirical investigations of blended AI-human instructional models remain relatively limited, particularly in discipline-specific English for Specific Purposes (ESP) contexts. This limitation highlights the need for further research examining how AI can be systematically integrated with teacher instruction to improve multiple dimensions of writing performance rather than focusing exclusively on linguistic accuracy

2.5 Ethics, Privacy, and Long-Term Implications

In addition to the pedagogical benefits of AI, its implementation in education raises important ethical, privacy, and academic integrity concerns. Scholars have emphasized that issues such as data protection, responsible use of AI-generated content, authorship, transparency, and academic honesty need to be carefully considered when incorporating generative AI into educational practices (Holmes et al., 2021; Dwivedi et al., 2023). These concerns highlight the importance of developing AI literacy among both educators and students so that users are able to understand the capabilities and limitations of AI systems and employ them responsibly within academic contexts.

Despite the growing conceptual discussion surrounding these issues, empirical evidence concerning the influence of ethical considerations on the actual adoption and use of AI in classrooms remains relatively limited. In particular, there is insufficient evidence regarding how students and teachers negotiate issues of privacy, authorship, reliability, and responsible use during AI-supported writing activities. This indicates that ethical considerations should not merely be treated as institutional or theoretical concerns but should also be examined as practical dimensions of AI integration in educational settings.



2.6 Critical Research Gaps and Rationale for the Current Study

Recent systematic reviews and meta-analyses indicate that research on AI-supported second-language (L2) writing has expanded rapidly but remains unevenly distributed across different dimensions of writing performance. Feng et al. (2025), for example, reported that more than 60% of empirical studies examining AI-assisted writing between 2020 and 2024 concentrated primarily on grammatical accuracy and lexical development. In contrast, fewer than 15% examined coherence, while only approximately 8% investigated clarity as a distinct writing construct. These patterns suggest that the current research landscape remains heavily oriented toward measurable, sentence-level linguistic outcomes, with considerably less attention devoted to higher-order textual qualities.

A similar pattern can be observed in Aryadoust's (2024) review of AI-based feedback research. Although more than 40 empirical studies have examined error correction and surface-level revision, only a relatively small proportion have investigated higher-order writing abilities, including idea development, organizational logic, rhetorical effectiveness, and clarity of expression. This imbalance is particularly important in ESP contexts, where successful communication requires learners not only to produce grammatically accurate sentences but also to organize discipline-specific information logically and communicate complex ideas clearly.

Methodological limitations also remain evident in the existing literature. A relatively small proportion of studies have employed quasi-experimental designs that directly compare AI-supported instruction with conventional teaching in authentic classroom environments. This limitation is especially relevant for non-English majors because their writing needs are closely connected to disciplinary contexts and professional communication requirements. Within ESP settings, students are expected to use English to communicate knowledge and ideas related to their particular fields; therefore, findings derived primarily from general English or English-major populations may not fully represent the instructional needs of non-English ESP learners.

Another important gap concerns the sustainability of AI-supported writing development. Existing research has predominantly examined immediate outcomes, whereas longitudinal studies and delayed post-test designs remain relatively uncommon. As a result, it remains unclear whether improvements in writing performance resulting from AI-assisted instruction are sustained over time or primarily reflect short-term effects associated with immediate automated feedback. Investigating sustained development is particularly important because the ultimate purpose of writing instruction is not simply to improve the quality of a single assignment but to foster transferable and independent writing competence.

The literature also reveals a need for greater empirical attention to the pedagogical and ethical dimensions of AI integration.



Although researchers have extensively discussed teacher mediation, responsible AI use, academic integrity, and potential overreliance on AI, these issues have often been addressed at a conceptual level rather than examined systematically within authentic classroom settings. Consequently, there remains limited empirical understanding of how AI-supported writing can be implemented while maintaining teacher involvement, student autonomy, academic integrity, and critical engagement with AI-generated feedback.

Method

This study utilized a quasi-experimental research design employing a pre-test–post-test control group structure to investigate the effects of ChatGPT-assisted instruction on students' academic writing performance. This design was selected because the research was conducted within an established university program in which students had already been assigned to intact classes through the institution's administrative scheduling system. Consequently, individual random assignment to experimental and control conditions was not feasible. The design is consistent with the quasi-experimental framework proposed by Creswell and Creswell (2018), which recognizes the use of naturally occurring or intact classes as an appropriate approach for educational research conducted in authentic instructional environments.

A true experimental design was considered inappropriate because randomization would have required reorganizing existing classes and potentially

disrupting the established academic timetable. Such changes could have affected the continuity of instruction and the normal learning environment. In addition, the study prioritized maintaining the integrity of the existing classes and minimizing disruption to students' regular learning activities. A mixed-methods approach was also considered; however, because the primary objective was to determine differences in writing outcomes between two instructional conditions, the quantitative quasi-experimental design was considered the most appropriate. Qualitative data obtained through student interviews were subsequently incorporated to provide contextual information concerning learners' experiences with the instructional intervention.

The absence of random assignment may potentially introduce selection bias. To minimize this threat to internal validity, several control measures were implemented. The two participating classes were comparable in terms of class size, academic characteristics, and English proficiency. Their initial English proficiency was established using institutional placement test scores administered at the beginning of the semester. Furthermore, both groups were taught by the same instructor and followed the same syllabus, learning objectives, writing assignments, assessment criteria, instructional duration, and classroom schedule. Maintaining these conditions helped reduce potential differences associated with teacher characteristics, curriculum content, instructional time, and learning environments.



The pre-test was also administered before the intervention to establish students' baseline writing performance and determine the comparability of the two groups prior to treatment. By examining students' initial performance, the study was able to account for potential pre-existing differences when interpreting post-test results. Thus, the design provided a reasonable level of control over alternative explanations and increased the likelihood that differences observed after the intervention could be attributed to the instructional treatment.

Within this research framework, the experimental group participated in writing instruction that incorporated ChatGPT as a supervised and guided writing assistant, whereas the control group received conventional teacher-led writing instruction without AI-based assistance. The comparison between these two conditions enabled the researcher to examine the relative effectiveness of ChatGPT-assisted and traditional approaches while preserving the authentic characteristics of an ESP classroom in a higher education setting.

The participants consisted of 60 undergraduate students enrolled in the Visual Communication Design program at a public university in Indonesia. A convenience sampling technique was employed because the study relied on intact classes that had already been established through the university's academic scheduling system. This sampling strategy allowed the research to be conducted within the existing instructional structure without requiring changes to students' class assignments or regular academic activities.

Two parallel classes were selected based on predetermined inclusion criteria. Participants were required to be non-English-major students enrolled in a compulsory English for Specific Purposes (ESP) writing course and to have completed the university's English placement test at the beginning of the semester. Students who were absent during either the pre-test or post-test were excluded from the final dataset to maintain the completeness and consistency of the data used for analysis.

The two classes were broadly comparable in demographic and academic characteristics. Most participants were first- and second-year undergraduate students between 18 and 21 years of age, and the gender composition of the two classes was generally balanced in accordance with the enrollment characteristics of the Visual Communication Design program. Their English proficiency levels were also considered comparable based on the results of the standardized placement test administered by the university at the beginning of the semester.

Prior to data collection, the researcher met with the students during regular class sessions to explain the purpose and procedures of the research. Information regarding the potential benefits and risks of participation, the voluntary nature of involvement, and participants' rights was clearly communicated. Written informed consent was obtained from all participants before the study commenced. Students were informed that participation was voluntary and that they had the right to withdraw from the



study at any point without experiencing academic consequences.

Participant confidentiality and anonymity were maintained throughout the research process. Numerical identification codes were assigned to students' writing samples, survey responses, and interview transcripts so that personal identities could not be directly associated with the collected data. Access to the coded dataset was restricted to the research team, and the electronic files were stored securely using encrypted storage. The research procedures received approval from the university's Institutional Review Board (IRB), confirming that the study complied with relevant ethical requirements for research involving human participants. These procedures were implemented to ensure that participant recruitment, data collection, storage, and analysis were conducted in accordance with established ethical research principles.

The research was conducted over a nine-week period using a structured procedure designed to maintain consistency between the experimental and control groups while enabling the systematic implementation of ChatGPT-assisted writing instruction. During the first week, students in both groups completed a pre-test designed to assess three dimensions of writing performance: clarity, coherence, and accuracy. The pre-test established baseline performance and provided an initial basis for examining the comparability of the two groups before the intervention.

The instructional intervention was implemented during Weeks 2 through 6. Both

classes met twice a week, with each session lasting 100 minutes. The experimental group received writing instruction in which ChatGPT was incorporated as a guided writing support tool. Before the intervention began, students participated in a 30-minute orientation session conducted by the instructor. The orientation introduced students to the basic functions of ChatGPT and demonstrated responsible strategies for using the tool during the writing process. Students were shown how to construct appropriate prompts, request feedback on specific aspects of writing, and evaluate and revise their drafts based on AI-generated suggestions.

Examples of prompts introduced during the orientation included requests such as "Revise this paragraph to improve coherence," "Provide clearer topic sentences for this outline," and "Suggest corrections for grammatical accuracy while keeping my tone unchanged." These examples were intended to demonstrate how students could use ChatGPT for specific writing purposes rather than relying on the system to generate complete texts independently.

During the intervention sessions, students first developed and drafted their own written work before consulting ChatGPT for selected aspects of the writing process. The AI tool was used to support idea development, paragraph organization, textual coherence, and surface-level linguistic revision. Students were encouraged to critically examine the feedback generated by ChatGPT and determine whether particular suggestions were appropriate for their intended meaning and writing context. The



instructor continuously monitored students' interactions with the AI system and provided clarification when necessary. This supervision was intended to prevent students from accepting AI-generated suggestions without critical evaluation.

To minimize the influence of external variables, the use of ChatGPT was restricted to supervised classroom sessions. Students were not permitted to use the tool outside the designated instructional activities, and revisions involving ChatGPT were completed during monitored sessions. This arrangement helped ensure that the intensity and nature of AI exposure were relatively consistent across participants in the experimental group.

The control group participated in the same syllabus, writing topics, learning activities, and assignment sequence as the experimental group but received conventional teacher-mediated instruction without AI assistance. The instructor provided written and oral feedback addressing the same primary writing dimensions considered in the experimental condition, namely clarity, coherence, organization, and accuracy. Both groups completed an equivalent number of writing assignments and received the same amount of instructional time, thereby reducing potential differences associated with unequal exposure to learning activities.

At the end of the instructional intervention, students in both groups completed a post-test during Week 7. The post-test was designed to be comparable in difficulty and task requirements to the pre-test and was evaluated using the same

analytic writing rubric. The use of equivalent assessment procedures allowed changes in students' writing performance to be examined systematically across the two instructional conditions.

Weeks 8 and 9 were devoted to collecting qualitative information concerning students' experiences with the learning intervention. Participants completed perception surveys and selected students participated in semi-structured interviews. The qualitative component focused particularly on students' perceptions of confidence, autonomy, and their experiences during the writing process. The combination of performance-based assessment and learner perception data enabled the study to examine not only whether students' writing performance changed following the intervention but also how students experienced and perceived the use of ChatGPT as part of their writing instruction.

Findings And Discussion

To evaluate whether students within each group made significant progress from the pre-test to the post-test, a paired-sample t-test was administered. A visual summary of these shifts is illustrated in Figure 1.

- Control Group: The mean writing score for this group progressed from 76.53 on the pre-test to 80.56 on the post-test. The paired-sample t-test confirmed that this advancement was statistically significant, $t(31) = -7.82, p < .001$. This indicates that conventional teaching methods still yielded measurable improvements in student writing performance.



- **Experimental Group:** For the group receiving AI-assisted instruction, the mean score climbed from 76.67 (pre-test) to 85.14 (post-test). The paired-sample t-test verified a highly significant improvement, $t(29) = 0.0001$, $p < .001$. These metrics demonstrate that integrating ChatGPT into the instruction generated more substantial academic gains than the traditional approach.

An independent-sample t-test was performed on the post-test scores to determine the variance in improvement across the two conditions. The analysis indicated a statistically significant difference between the final scores of the control group ($M = 80.56$) and the experimental group ($M = 85.14$), $t(58) = 0.021$, $p < .05$. This finding confirms that students who utilized ChatGPT significantly outperformed their conventionally taught peers, particularly in the domains of structural clarity and coherence.

Ultimately, while both pedagogical methods successfully advanced the students' writing skills, the inclusion of ChatGPT exerted a more pronounced positive effect. This outcome corroborates prior literature asserting that AI-driven feedback can accelerate learning trajectories and cultivate learner independence (Kasneci et al., 2023; Wang et al., 2023).

The questionnaire data indicated that most participants held favorable attitudes toward incorporating AI tools into writing coursework. Specifically, 78% of the respondents agreed that ChatGPT refined their grammatical precision and textual

structure, while 64% noted that the tool incentivized them to engage in more frequent draft revisions.

Beyond cognitive developments, affective benefits were also observed: 59% of the students reported lower levels of writing anxiety alongside heightened confidence when collaborating with the AI tool. Nonetheless, certain apprehensions persisted. Roughly 21% of the students questioned the reliability of the AI's feedback, and 17% expressed concerns regarding a potential overreliance on automated text generation. These insights suggest that while AI enhances both the cognitive and emotional facets of writing, its application requires careful teacher mediation to maximize pedagogical value. A comprehensive breakdown of these survey metrics is detailed in Table 3.

A thematic analysis of the interview transcripts offered deep insights into the students' lived experiences using ChatGPT during their writing tasks. Three predominant themes emerged from the data, detailing shifts in student confidence, autonomy, and their overall critical perceptions of the tool.

Table 4. Themes from Interview Analysis

Theme	Description	Sample Quote
Enhanced Confidence	Students felt less anxious and more willing to take risks in writing.	"I feel comfortable experimenting"
Increased Autonomy	Learners became more responsible for revising	"ChatGPT made me more responsible"
Mixed Reliability	Some concerns about the accuracy and contextual	"Sometimes the suggestions didn't always make sense"



The outcomes of this investigation indicate that integrating ChatGPT into instruction yields substantially greater advancements in writing clarity, coherence, and accuracy compared to conventional ESP pedagogical approaches. While these observations corroborate the expanding body of literature validating generative AI as a tool for linguistic enhancement (Kasneci et al., 2023; Generative AI-Assisted Feedback, 2025), this study expands upon current research by demonstrating that the most pronounced progress occurred in textual coherence—a domain where earlier investigations have remained inconclusive or overly conservative. Notably, multiple scholars have argued that AI's deficits in contextual and rhetorical comprehension restrict its utility regarding higher-order text characteristics (Liu & Gu, 2023; Chan & Tsi, 2023). The data gathered here partially dispute this stance; under structured and guided conditions, students leveraged ChatGPT not merely for surface-level mechanical edits but for restructuring the conceptual flow of their ideas. This suggests that the efficacy of AI on global text quality depends heavily on pedagogical mediation rather than the intrinsic constraints of the technology itself. This specific nuance is generally overlooked in contemporary literature and emphasizes the necessity for a more conditional evaluation of AI capabilities.

Furthermore, the pre-test to post-test developmental trajectories challenge the implicit assumption within much of the AI-based writing literature that such technologies uniformly benefit all learners.

Subgroup analyses demonstrated disproportionate progress among lower-proficiency students, indicating that AI can function as an effective compensatory scaffold. Consequently, this finding complicates previous assertions that AI exacerbates educational disparities by disproportionately favoring already-skilled writers (cf. critiques by Erito, 2023). In this instance, the opposite dynamic emerged, demonstrating that student gains depend not solely on baseline proficiency, but on whether the AI feedback is embedded within a structured, process-oriented writing environment. This outcome aligns directly with process writing theory, which attributes skill acquisition to recursive revision cycles; however, the deployment of AI intensified these cycles by drastically reducing the latency period for feedback—a specific mechanism that has rarely been operationalized in previous studies.

The qualitative insights shed further light on the tensions underlying the second research question (RQ2). Participants reported heightened levels of confidence and autonomy, aligning with prior studies that emphasize the affective benefits of AI tools (Wang et al., 2023; Apriani et al., 2024). Nevertheless, their skepticism regarding the precision, relevance, and disciplinary appropriateness of certain AI outputs reflects the early stages of feedback literacy development (Carless & Winstone, 2023). This qualification counters overly optimistic depictions of AI as an inherently empowering agent. Instead, these findings show that learner autonomy expands only when students learn to critically evaluate AI



feedback rather than passively accept it. This distinction is rarely articulated in current AI-in-education discourse, yet it remains crucial for preventing exaggerated claims regarding learner agency.

Conclusion

This study demonstrates that while both traditional and AI-assisted instructional frameworks successfully improved student writing performance, the experimental group utilizing ChatGPT achieved more substantial progress than the control group. The empirical data confirm that AI-driven feedback can optimize structural clarity, coherence, and overall textual quality by providing immediate, individualized support. Furthermore, students within the experimental condition displayed enhanced writing confidence and greater autonomy, indicating that ChatGPT can positively shape both skill acquisition and learner attitudes.

However, these insights simultaneously underscore the irreplaceable role of teacher guidance. While automated feedback is highly efficient, it frequently lacks the contextual and rhetorical precision required for discipline-specific academic writing. Consequently, ChatGPT should be conceptualized as a pedagogical complement to, rather than a replacement for, conventional classroom instruction. Future research should investigate the long-term durability of AI-assisted writing skills, its utility within highly specialized academic disciplines, and the broader ethical implications governing its implementation in higher education. Such investigations will be vital to optimizing the integration of artificial

intelligence into sustainable educational practices.

References

- Aryadoust, V. (2024). *A systematic review of AI-based automated written feedback research*. *ReCALL*, 36(1), 1–23.
<https://doi.org/10.1017/s0958344023000265>
- Apriani, E., Cardoso, L. M., Obaid, A. J., Muthmainnah, & Dewi, R. (2024). Impact of AI-powered chatbots on EFL students' writing skills, self-efficacy, and self-regulation: A mixed-methods study. *Global Educational Research Review*, 1(2), 57–72.
- Azennoud, A. (2024). Enhancing writing accuracy and complexity through AI-assisted tools among Moroccan EFL university learners. *International Journal of Linguistics and Translation Studies*, 5(4), 211–226.
<https://www.ijlts.org/index.php/ijlts/article/view/529>
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
<https://doi.org/10.4324/9780203832400>
- Chan, C. K. Y., & Tsi, L. H. Y. (2023). *The AI revolution in education: Will AI replace or assist teachers in higher education?* arXiv.



- <https://doi.org/10.48550/arXiv.2305.01185>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Datskiv, O., Zadorozhna, N., & Shon, P. (2024). Developing future teachers' academic writing and critical thinking skills using ChatGPT. *International Journal of Emerging Technologies in Learning (iJET)*, 19(8), 4–20. <https://doi.org/10.3991/ijet.v19i08.49935>
- Discover Education. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4, Article 170. <https://doi.org/10.1007/s44217-025-00602-7>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Erito, S. N. P. (2023). Exploring ESP students' perception toward the potential of artificial intelligence to promote students' self-efficacy in English writing skill. *Journal of English Language Learning*, 7(2). <https://ejournal.unma.ac.id/index.php/jell/article/view/7598>
- Feng, H., Li, K., & Zhang, L. J. (2025). *What does AI bring to second language writing? A systematic review (2014–2024)*. Language Learning & Technology. <https://doi.org/10.64152/10125/7362>
- Ferris, D. R. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, 19, 6–23. <https://doi.org/10.1016/j.asw.2013.09.004>
- Fitriani, E., & Fakhruddin, A. (2024). Exploring students' experience in using AI to assist their writing. *Journal of English Language Learning*, 8(1). <https://ejournal.unma.ac.id/index.php/jell/article/view/10028>
- Harahap, P. H. K., Arian, R., Harahap, R., Rahmadani, S., & Audina, F. (2023). The influence of reading habits, critical thinking ability, and mastery of syntactic structures on scientific writing ability. *Jurnal Pendidikan Tambusai*, 7(3), 30960–30968. <https://jptam.org/index.php/jptam/article/view/12040>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O.



- C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). *Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Horwitz, E. K. (2016). Factor structure of the Foreign Language Classroom Anxiety Scale: Theoretical advances and practical applications. *TESOL Quarterly*, 50(4), 932–937. <https://doi.org/10.1002/tesq.295>
- Hyland, K., & Hyland, F. (2019). Feedback in second language writing: Contexts and issues. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (2nd ed., pp. 1–19). Cambridge University Press. <https://doi.org/10.1017/9781108635547>
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kukulska-Hulme, A. (2021). Mobile-assisted language learning [Revised and updated]. In M. Thomas & H. Reinders (Eds.), *Contemporary computer-assisted language learning* (pp. 219–236). Bloomsbury.
- Lee, I. (2017). Classroom writing assessment and feedback in L2 school contexts. *Springer*. <https://doi.org/10.1007/978-981-10-3924-9>
- Pratama, A., & Sulistiyo, U. (2024). A systematic review of artificial intelligence in enhancing English foreign learners' writing skill. *PPSDP International Journal of Education*, 3(2), 170–181. <https://ejournal.ppsdp.org/index.php/pijed/article/view/299>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151, 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- Supardan, D. (2024). *Impact of AI-powered ChatBots on EFL students' writing skills, self-efficacy, and self-regulation: A mixed-methods study*. *Global Educational Research Review*, 1(2), 57–72. <https://doi.org/10.71380/GERR-08-2024-8>
- Supporting the Writing Process with Technology for Students with*



- Disabilities. (2024). *Journal of Language and Education*, 3(3), 79-94.
<https://doi.org/10.17323/jle.2024.18708>
- Tanjung, S., Nasution, N., & Gusmaneli. (2025). The application of direct instruction strategy in shaping students' critical thinking competence. *Jurnal Arjuna: Publikasi Ilmu Pendidikan, Bahasa dan Matematika*, 3(2).
<https://journal.aripi.or.id/index.php/Arjuna/article/view/1761>
- Tanjung, S., Nasution, N., & Gusmaneli. (2025). The application of direct instruction strategy in shaping students' critical thinking competence. *Jurnal Arjuna: Publikasi Ilmu Pendidikan, Bahasa dan Matematika*, 3(2).
<https://journal.aripi.or.id/index.php/Arjuna/article/view/1761>
- Wang, Y., Liang, J. C., & Tsai, C. C. (2023). Students' perspectives on using artificial intelligence writing assistants in higher education. *Computers & Education*, 195, 104695.
<https://doi.org/10.1016/j.compedu.2023.104695>
- Williyan, A., Fitriati, S. W., Pratama, H., & colleagues. (2024). *AI as co-creator: Exploring Indonesian EFL teachers' collaboration with AI in content development*. Teaching English with Technology.
<https://doi.org/10.56297/vaca6841/lrdx3699/rzoh5366>
- Yu, H., & Xie, Q. (2025). *Generative AI vs. Teachers: Feedback Quality, Feedback Uptake, and Revision*. *Research Quarterly in Language Teaching*, 47(1), 113–137.
<https://doi.org/10.32038/ltrq.2025.47.07>
- Yuliana, I., Hestiningtyas, Y. P., & Sundaygara, C. (2020). Analysis of critical thinking skills in terms of students' argumentation abilities through the Argument Based Science Inquiry (ABSI) model. *Jurnal Riset dan Kajian Pendidikan Fisika*, 7(2).
<https://journal.uad.ac.id/index.php/JRKPF/article/view/17093>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(39).
<https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X. (2023). ChatGPT for next generation science learning. *Science & Education*, 32, 423–430.
<https://doi.org/10.1007/s11191-023-00390-5>
- Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2



ELECTRUM

The English Language and Education Spectrum

P-ISSN : 2746-4717

E-ISSN : 2808-2818

writing. *Assessing Writing*, 36, 90–
102.

[https://doi.org/10.1016/j.asw.2018.
02.004](https://doi.org/10.1016/j.asw.2018.02.004)